

# IBM Data Science Professional Certification by Coursera

## Applied Data Science Capstone

### Opening a New Multiplex Business in Hyderabad, India

Phalgun Reddy Bobbili

August 2020



## **Introduction**

The Indian film industry is the produces the greatest number of movies every year than any other film industry. Every year, around 2000 movies are released across many languages in the country. In 2015, the total box office gross was over \$2.1 billion, third highest in the world. The people of India love going to the cinemas to enjoy the experience of a big screen viewing. Even when the OTT platforms are rising and gaining users and traction in the country, the box office collections have not decreased. This is because the public still prefer to go to the cinema theatres for the that big screen experience. This is even more applicable to the southern part of the country, where the love for cinemas even more special. Hyderabad is one of the most developed metropolitan cities in the country and located in southern India.

There are basically two kinds of cinema theatres. The single screen theatres and the multiplexes which offer multiple screens. The pubic these prefer going to multiplexes over the single screen theatres. The multiplexes come with a various shopping opportunities and food courts. Also, the multiplexes offer multiple movies at many different times of the day, allowing the people to choose their preferred movie at their preferred time. These factors enable the people to plan a whole evening at one location, attracting a lot of customers. On the other hand, multiplexes require huge investments to be established. Hence, a lot of planning should go into it. One of the most important aspects of a multiplex that drive its success is the location of the multiplex. Factors like the number of competitors in the area, the demographics of the area, etc. influence the success of the business heavily.

## **Business Problem**

The objective of this capstone project is to analyze the city of Hyderabad and try to select the most suitable locations to open a new multiplex business in the city. This project aims to achieve this by using data science methodologies and machine learning techniques like clustering.

## **Data**

To successfully finish this project, we will need the following data:

- We will be analyzing the various locations of the city by neighborhoods. Hence, we will need a list of neighborhoods in Hyderabad.
- We will use the coordinates of the neighborhoods to plot Folium maps of the city. We will need the latitudes and the longitudes of the neighborhoods for this.
- We will retrieve venues data from Foursquare API and extract the multiplexes out of the venues data. Venue data, particularly data related to Multiplexes. This data will be used to cluster the neighborhoods.